SUPPLEMENTARY MATERIAL

METHODS

Study population. Patients eligible to sarilumab were required to have confirmed SARS-CoV-2 infection by reverse-transcriptase polymerase-chain-reaction on nasal-pharyngeal swab and radiologically documented bilateral pneumonia. In addition patients were required to have severe COVID-19 as defined by either ≤ 92% of oxygen saturation while breathing ambient air or by a partial pressure of arterial oxygen/fraction of inspired oxygen (PaO_2/FiO_2) ratio ≤ 300 mmHg on supplemental oxygen, and a hyper-inflamed phenotype as defined by an elevation of lactate dehydrogenase (LDH) above the upper limit of normal (ULN), and by at least one of the following: C-reactive protein (CRP) ≥ 100 mg/L; IL-6 ≥ 40 pg/ml; or ferritin (≥ 900 ng/ml). Patients hospitalized for more than 14 days, on concomitant or previous immunosuppressive agents, or mechanically ventilated were excluded. Patients with uncontrolled systemic infections, total neutrophil count < 1500/mm^3, serum levels of alanine aminotransferase (ALT) and aspartate aminotransferase (AST) more than five times the ULN, diverticulitis/diverticulosis, and pregnant women were also excluded.

Standard of care treatment for Covid-19 patients hospitalized at San Raffaele Hospital. *Antiviral therapy*: in the absence of specific contraindications, all patient included in this study received: Hydroxychloroquine 200 mg BD orally and Lopinavir/Ritonavir, 400/100 mg BD orally. *Steroids*: glucocorticoids were not part of the standard of care and were not used to treat enrolled patients. Being on other immunosuppressive agents was considered an exclusion criteria. *Antibiotic therapy*: all patients received an initial empiric antibiotic coverage for community acquired/hospital acquired pneumonia based on either daily intravenous Ceftriaxone

Della-Torre E, *et al. Ann Rheum Dis* 2020;0:1–9. doi: 10.1136/annrheumdis-2020-218122

2 g or Azithromycin 500 mg. For patients with negative cultures and decreasing inflammatory markers ceftriaxone was discontinued after 6 days of therapy.

Computed tomography scan protocol and analysis. Chest CT scans were obtained using 64-channel CT scanner (GE LightSpeed VCT Cardiac Pro) during a full-inspiration breath-hold using the following parameters: detector configuration of 64 ×0.625 mm, tube voltage of 120 kVp, tube current of 180–200 mAs and pitch of 1.375. Image data sets were reconstructed with 1-2 mm slice thickness range using both sharp kernels (B70f) with a standard lung window (1500 width; -500 center). All chest CT scans were analysed with a pulmonary-dedicated software (CT COPD, IntelliSpace Portal 8.0, Philips Healthcare, Cleveland, OH, USA). Two expert radiologists blinded to clinical data performed all image analyses. CT scans were analysed with a pulmonary-dedicated software (CT COPD, IntelliSpace Portal 8.0, Philips Healthcare, Cleveland, OH, USA) for automatic lung segmentation and derivation of volumes involved by the inflammatory process. A fully automatic lung volume segmentation and analysis of lung parenchyma histogram was obtained; in case of inadequate segmentation, the users adjusted lung contours with a manual tool. Attenuation of each voxel within segmented lungs was computed automatically. Three thresholds of Hounsfield units (HU) were established (cit. 1 e 2): between -950 and -740 as "normal parenchyma"; between -740 and -350 as "intermediate density"/"ground glass", and over -350 as "consolidation". All volumes were calculated and expressed as percentage and volume.

Statistical analysis. Statistical analysis was performed using Prism software 8.0 (GraphPad Software, La Jolla, CA, USA). Continuous variables are reported as medians and interquartile ranges. Categorical variables are reported as numbers and percentages. Wilcoxon rank-sum tests

Della-Torre E, *et al. Ann Rheum Dis* 2020;0:1–9. doi: 10.1136/annrheumdis-2020-218122

were applied to continuous variables and two-tailed Fisher's exact tests were used for categorical variables. Survival analysis was performed with the Kaplan-Meier approach, log-rank test was used to compare survival curves. Survival and clinical and laboratory features at baseline were analysed using proportional hazard Cox regression models. Results of the Cox regression model are presented as hazard ratio (HR) with 95% confidence interval. Receiver operating characteristic (ROC) curves were generated based on the assumption that ascending rank of PaO_2/FiO_2 ratio and lung consolidation would increase the likelihood of reaching or not reaching clinical improvement, respectively. Linear correlations were measured by Spearman's correlation coefficient. P-values <0.05 were considered statistically significant. Values are presented as median and interquartile range (IQR), unless specified otherwise.

Della-Torre E, *et al. Ann Rheum Dis* 2020;0:1–9. doi: 10.1136/annrheumdis-2020-218122

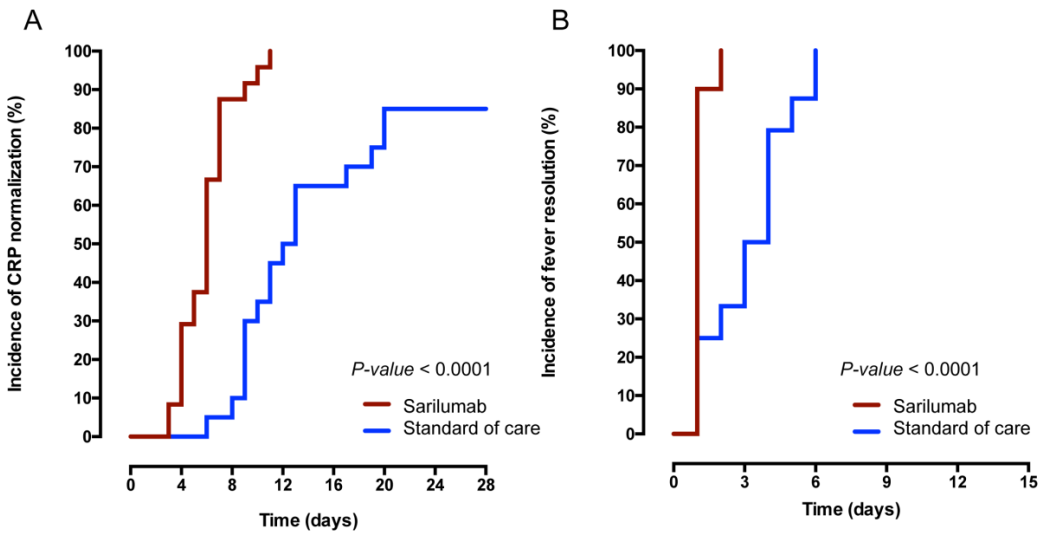
TABLE S1.

	Improved (n=18)	Not-improved (n=10)	p-value	Survived (n=23)	Dead (n=5)	p-value
Age, yr	57 (52-60)	56 (49-67)	0.87	56 (52-59)	67 (51-72)	**0.02**
Sex (male/female)	14/4	6/4	0.41	16/7	4/1	0.99
Hospitalization before enrolment, days	2 (1-3)	3 (2-4)	0.07	2 (1-3)	3 (2-3)	0.19
Non-invasive mechanical ventilation, n (%)	10 (55)	9 (90)	0.1	13 (56)	5 (100)	0.56
Coexisting conditions, n (%)						
Arterial hypertension	9	2	0.22	10	1	0.61
Tobacco smoking	3	2	0.99	5	0	0.55
Dislipidemia	3	1	0.99	3	1	0.99
Coronary artery disease	3	1	0.99	3	1	0.99
Type 2 diabetes	3	3	0.63	5	1	0.99
Chronic obstructive pulmonary disease	0	1	0.35	1	0	0.99
Chronic renal failure	1	1	0.35	1	1	0.33
PaO_2/FiO_2 ratio (mmHg)	135 (89-175)	80 (64-131)	**0.02**	122 (89-172)	77 (62-121)	**0.02**
Axillary Temperature, °C	38 (37-38.9)	38 (37-39)	0.89	38 (37-38.9)	37 (36.7-38.5)	0.28
Laboratory values						
CRP, (normal < 6 mg/L)	152 (116-202)	163 (123-259)	0.53	153 (122-211)	139 (120-220)	0.81
Ferritin, (normal 30-400 ng/mL)	1376 (1056-3200)	1542 (1153-3542)	0.77	1269 (1056-3200)	1046 (827-1178)	0.12
LDH, (normal 125-220 IU/L)	469 (367-589)	510 (397-682)	0.64	479 (389-632)	567 (305-732)	0.67
IL-6, (normal <7 pg/mL)	39 (27-96)	75 (57-258)	0.06	46 (24-117)	68 (24-101)	0.92
AST, (normal 5-35 IU/L)	55 (32-91)	48 (23-153)	0.99	58 (27-105)	51 (31-147)	0.99
ALT, (normal 6-59 IU/L)	49 (26-82)	49 (19-120)	0.99	51 (27-90)	62 (22-163)	0.79
Creatine kinase (normal 20-195 IU/L)	101 (70-238)	47 (18-155)	0.34	87 (29-162)	87 (22-161)	0.99
D-dimer (normal 0.27-0.77 μg/mL)	1.91 (0.81-9.47)	2 (1.25-6.51)	0.94	1.9 (0.81-4.99)	2.36 (1.72-3)	0.79
Platelets (normal 130-400 cells/μL)	221 (192-380)	299 (144-434)	0.83	227 (166-410)	407 (142-453)	0.72
Lymphocytes (normal 1000-4800 cells/μL)	700 (450-1100)	800 (775-1500)	0.23	800 (575-1225)	1000 (550-1250)	0.84
Neutrophils (normal 1800-7000 cells/μL)	4800 (3975-7600)	7450 (4200-9275)	0.41	5100 (4000-8600)	7700 (4800-9500)	0.31
Radiological features*						
Consolidation (cc)	318 (180-546)	993 (703-1284)	**0.03**	318 (180-546)	993 (703-1284)	**0.03**
Consolidation (%)	7.2 (4.2-16.9)	34.3 (14.4-54.3)	0.08	7.2 (4.2-16.9)	34.3 (14.4-54.3)	0.08
Ground glass (cc)	955 (787-982)	2138 (1070-3207)	0.06	955 (787-982)	2138 (1070-3207)	0.06
Ground glass (cc)	21.3 (17.3-29.9)	55.6 (45.3-65.9)	0.06	21.3 (17.3-29.9)	55.6 (45.3-65.9)	0.06

Unaffected lung (cc)	3045 (1861-3282)	484 (10-958)	0.06	3045 (1861-3282)	484 (10-958)	0.06
Unaffected lung (%)	73.7 (54.1-75.4)	10.1 (1-19.7)	0.06	73.7 (54.1-75.4)	10.1 (1-19.7)	0.06

Baseline univariate predictors of clinical improvement and survival in patients treated with standard of care. Variables associated with clinical improvement (defined as live discharge from hospital or 2 points improvement from baseline on the six-category ordinal scale) and with overall survival at 28 days in patients treated with standard of care were studied by univariate analysis. Continuous data are reported as median (IQR). Categorical data are reported as number of patients (n) and percentage (%). A p-value < 0.05 was considered statistically significant. **Abbreviations**: partial pressure of arterial oxygen/fraction of inspired oxygen ratio (PaO2/FiO2 ratio); Lactate dehydrogenase (LDH); C-reactive protein (CRP); alanine aminotransferase (ALT); aspartate aminotransferase (AST). *Computed tomography scan was performed at baseline in n = 8 patients treated with standard of care. Because four of them experienced clinical improvement at 28 days follow-up and four deceased, the values of CT scan findings in patients improved or survived and in patients not improved or dead overlap.

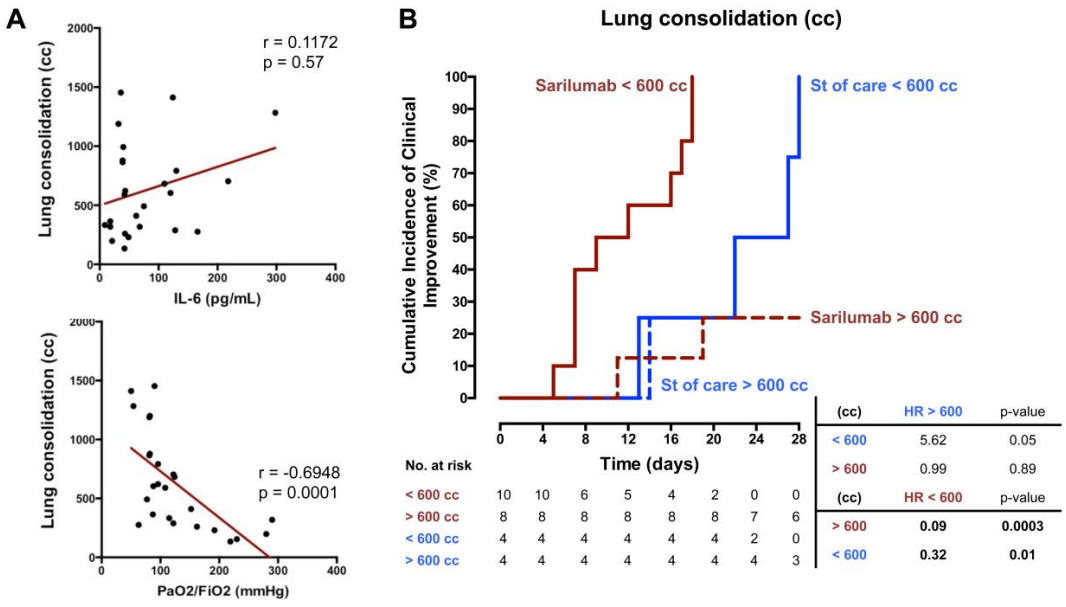
FIGURE S1



Inflammatory markers. Cumulative incidence of CRP improvement (**A**), and fever normalization (**B**) at 28 days are outlined in Kaplan-Meier curves. **Abbreviations**: C-reactive protein (CRP).

Della-Torre E, *et al. Ann Rheum Dis* 2020;0:1–9. doi: 10.1136/annrheumdis-2020-218122

FIGURE S2.



Predictors of clinical improvement in patients treated with sarilumab. (A) Correlation studies between predictors of clinical improvement identified by univariate analysis in patients treated with sarilumab. **(B)** Cumulative incidence of clinical improvement in patients treated with sarilumab and standard of care defined as discharge from hospital or 2 points improvement from baseline on the six-category ordinal scale according to baseline total volume of lung consolidation**.** Clinical improvement was observed in 100% of cases with < 600 cc of consolidated lung and in only 25% of cases with lung consolidation > 600 cc. Median time to clinical improvement was shorter in patients with < 600 cc of consolidated lung (11 (7-17) days vs 28 (21-28); HR 0,09; 95% CI 0,03-0,35; p-value = 0,0003). **Abbreviations:** hazard ratio (HR).